Dynamic Facial Caricaturing System Based on the Gaze Direction of Gallery

Kazuhito Murakami
Aichi Prefectural University
Nagakute-cho, Aichi 480-1198, Japan
murakami@ist.aichi-pu.ac.jp

Masafumi Tominaga and Hiroyasu Koshimizu
SCCS, Chukyo University
Kaizu-cho, Toyota 470-0393, Japan
{tomy, hiroyasu}@koshi-lab.sccs.chukyo-u.ac.jp

Abstract

Facial caricaturing is a representation process of human visual impression onto a paper or other media. Facial caricaturing should be discussed from multiple viewpoints of three relations among the model, the caricaturist and the gallery. Furthermore, some kinds of interactive mechanism should be required between the caricaturist and the gallery.

In this paper, we propose a dynamic caricaturing system. In our system, the utilization of in-betweening method realizes the generation mechanism from the caricaturist to the gallery, and on the contrary, the utilization of eye-camera vision realizes the feedback mechanism from the gallery to the caricaturist. This is an original and unique point of our system. The gallery mounts an eye-camera on the head, and the system reflects visual characteristics of the gallery directly onto the works of facial caricature. After observing the image of the model and analyzing the gaze direction and distribution, the system deforms some characteristic and impressive facial parts more strongly than other non-impressive facial parts, and generates the caricature which is suited especially for the gallery. In this paper, we demonstrate experimentally the effectivity of this method to integrate these kinds of viewpoints.

1. Introduction

There are many types of media to present a facial caricature; (a)newspaper or magazine, (b)Web, (c)handmade drawing, and so on. In the case-(a), the caricaturists contribute his or her works to the editor, and some good works are selected by the editor and published. The gallery, that is the readers of the magazine, enjoy the caricature. It can be said that there is no interactivity in this mechanism, because the flow of information (interest) is one-way from the caricaturist to the gallery. In the case-(b) which uses Web [1], it is slightly but partially interactive in the sense that the gallery can select the model of the caricature. In the case-(c), a professional caricaturist draws a work with careful watching of the model and interesting conversation with some requests, "to be a little younger", "to be more beautiful", "to be handsome", and so on. So this types of caricaturing is interactive caricaturing.

Facial caricaturing is a representation process of human visual impression onto a paper or other media. From the viewpoint of computer vision, it is important to brush up the technique of drawing, and it is more important to clarify how the human vision extracts the feature points of the face and recognizes them in advance. In the conventional systems or researches concerning facial deformation or facial feature extraction [2, 3], caricaturing has been discussed in the viewpoint of the relation only between the model and the caricaturist, and the flow of information also has been treated as one-way from the caricaturist to the gallery. As the caricature varies according to who draws the caricature, the evaluation varies who observes it [4]. From these analytical considerations, facial caricaturing should be discussed from multiple viewpoints of these three relations among the model, the caricaturist and the gallery as shown in Fig.1. Furthermore, some kinds of interactive mechanism should be required between the caricaturist and the gallery.

In this paper, we propose an interactive caricaturing system. In our system, the utilization of in-betweening method realizes the generation mechanism from the caricaturist to the gallery, and on the contrary, the utilization of eye-camera vision realizes the feedback mechanism from the gallery to the caricaturist. This is an original and unique point of our system. The eye-camera is mounted on the head of the gallery, and visual characteristics of the gallery are directly reflected onto the works of facial caricature. After observing

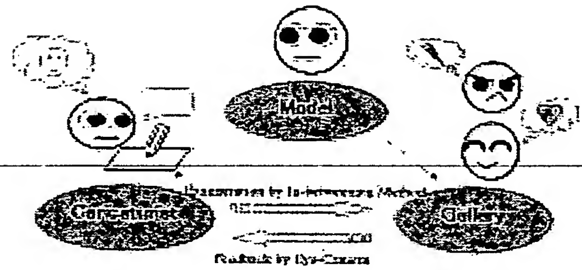


Figure 1. The relations among model, caricaturist and gallery.

the image of the model and analyzing the gaze distribution, the characteristic and impressive facial parts are more strongly deformed than other non-impressive facial parts, and finally the caricature which is suited especially for this gallery is generated.

In section 2, the basic formalism and the system configuration of our proposed mechanism are described. In section 3, some experimental results of gaze distribution by using eye camera vision are reported. In section 4, some examples of caricatures are presented by using our interactive system.

2. System Configuration

2.1. Basic Formalism of Caricature Generation

Facial caricaturing should be discussed from all the view points of the model, the caricaturist and the gallery as described in section 1. As shown in Fig.1, the caricature generation is described as

$$Q = func_C(P, S_1, K) \quad (1)$$

$$K = func_G(P, S_2), \quad (2)$$

where, P is input face, S_1 and S_2 are reference, or standard, faces (afterward, mean face is used), K is visual interest parameter of a gallery, and Q is caricature of the person P. Subscripts C and G mean caricaturist and gallery, respectively.

Suppose that Eq.(1) could be expanded and described in the following polynomial expression as

$$Q = func_C(P, S) + func_G(P, S) + func_C(func_G(P, S)) + ... \quad (3)$$

The second term in Eq.(3) is a feedback element from a gallery. The third term and its successors, the recursive feedback elements, are omitted hereafter in this paper, because these values are relatively not so serious.

Next, we assume that $func_C$ and $func_G$ could be characterized by the individuality feature and its weight. As one of the most simple expressions, let the individuality feature be the deviation $P - S$ and its weight be b, then caricature Q can be simply described as

$$Q = P + b_C(P - S) + b_G(P - S) \quad (4)$$

$$S = average(P_1, P_2, ..., P_M). \quad (5)$$

S is the mean face of the persons $P_1 \sim P_M$. This mean face is often used to extract individuality features. Of course, the above formalism is not a unique representation. Since the caricature depends on the caricaturist and the gallery, the functions or values b_C and b_G can be changed according to the principle of the system.

One of the main subjects is to introduce the gallery factor b_G. In our system, the degree of distributive interest for each facial feature point of the model P is measured by using eye camera, and this value is utilized as the factor of gallery b_G. So, in our system, caricature generation could be simply formalized as

$$Q = P + (b_C + b_G) \cdot (P - S). \quad (6)$$

2.2. Precise Description

In our system, facial features are represented by using N points of x and y coordinates of the facial contour as follows[2,9,10]:

$$x_i^{(Q)} = x_i^{(P)} + b_i \cdot (x_i^{(P)} - x_i^{(S)})$$
$$y_i^{(Q)} = y_i^{(P)} + b_i \cdot (y_i^{(P)} - y_i^{(S)}) \quad (7)$$
$$i = 1, 2, ..., N,$$

where, P is a primal input face, S is a reference face (mean face), Q is a deformed caricature and b is the degree of interest of the gallery. The mean face can be calculated by

$$x_i^{(S)} = \frac{1}{M}\sum_{j=1}^{M} x_i^{(P_j)}, \qquad y_i^{(S)} = \frac{1}{M}\sum_{j=1}^{M} y_i^{(P_j)} \quad (8)$$
$$i = 1, 2, ..., N,$$

where $x_i^{(P_j)}$ and $y_i^{(P_j)}$ are the x and y coordinates for the i-th feature point of the j-th normalized face data.

The degree of interest $b_i (i = 1, 2, ..., N)$ could be determined according to the distribution of gaze of the gallery. In our system, it is decided based on the distance between the gaze point (x_{gaze}, y_{gaze}) and the feature point (x_i, y_i) $(i = 1, 2, ..., N)$, and this is defined as the integration for m times measured data. After normalization, the degree of interest for the i-th point b_i is defined as

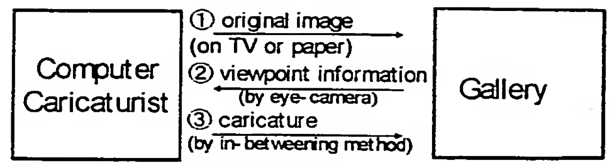


Figure 2. System configuration.

$$b_i = b_{offset} + \beta \cdot (\frac{d_{min}}{d_i})^{\alpha} \quad (9)$$

$$d_{min} = minimum(d_i)(i = 1, 2, ..., N) \quad (10)$$

$$d_i = \frac{1}{m}\sum_{t=1}^{m}((x_{gaze}^{(t)} - x_i)^2 + (y_{gaze}^{(t)} - y_i)^2)^{1/2} \quad (11)$$

here, b_{offset} is an offset value and decided experimentally. The value b_i of the most interested point becomes $b_{offset} + \beta$, and on the contrary, it becomes about b_{offset} around the uninterested points.

2.3. System Configuration

Our system configuration is simply summarized in Fig.2. In our system, the basic mechanism is composed of three sub-processes;
(1)presentation of original image,
(2)analysis of view point information (gaze distribution), and
(3)generation of caricature.

The basic idea to realize the feedback mechanism of sub-process (2) from a gallery to the caricaturist (=the system) is the utilization of eye-camera vision. This is an attempt to extract non-verbal visual information. Details of sub-processes (1) and (2) are described in the following section 3. The interactivity is realized by using this feedback mechanism.

3. Gaze Detection by Eye Camera and Analysis of ROI

Although some researchers have been trying to extract individuality features of the face by using image processing techniques, it is difficult to conclude that the measures represent individuality information sufficiently. In our research, we utilized the gaze information which is measured by using an eye camera mounted on the head.

Figure 3 shows a scene of extracting visual characteristic points by using an eye-camera. The direction of the eye can be calculated by using the difference of the reflection-ratio of white part and colored-part (iris and pupil). Since the gray image is also inputted through small CCD camera attached on the eye-camera, the system can calibrate where the gallery looks at in the gray image.

In order to extract visual characteristics of the gallery for the model, firstly, an original gray image of a person is displayed on the screen at the fixed place, here, the angle and the distance between the image and the gallery are calibrated in advance (about 1[m]) as shown in Fig.4. The size of the face is about 300–350[mm] (maximum size). Secondly, an instruction is given to the gallery to look at some characteristic point he/she feels impressive among the face. From the measured data obtained by eye-camera, the gaze distribution, fixation and saccade are calculated and analyzed.

In our system, the gaze distribution is analyzed by the following method. First,
(1)gaze data is sampled in every 16 [msec]'s, and
(2)the locus is calculated and the distribution of the region of interest (ROI) of the gallery is calculated.
In our system, totally 3.2[sec]'s data are used(during 5.0[sec]–8.2[sec]'s data after the display of the face is started).

Figure 5 shows an example of the examination. The model is Mr. Naoto Kan, a Japanese statesman. Figure 5(a) is an original image, and the gaze distributions of the left and right eyes are shown in Fig.5(b) and Fig.5(c), respectively. The horizontal and vertical axes show the location of the peak of the distribution. The result of distributions overlaid onto the original image with the marks L and R is shown in Fig.5(d). Although there is some differences between the left and right eye's distribution, it was clarified that the gallery

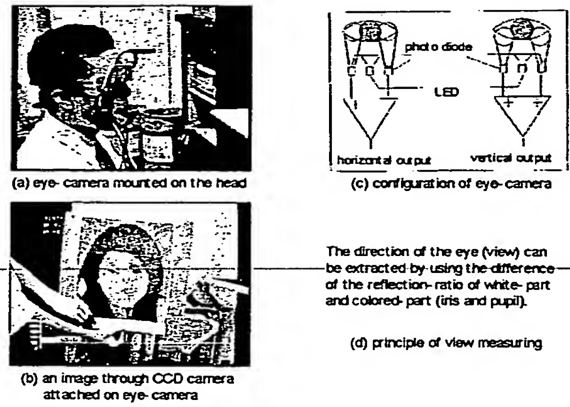


Figure 3. A scene of eye-camera mounted on the head and the gray image through eye-camera.

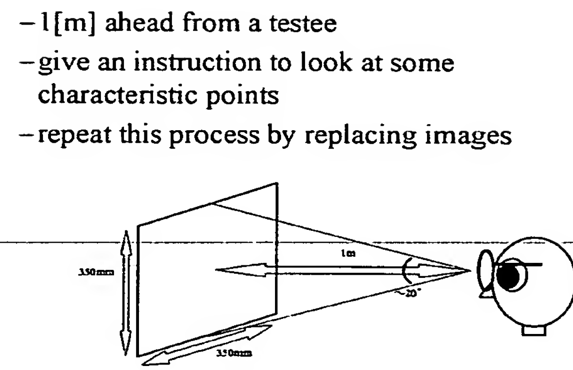


Figure 4. Presentation of original image.

looks mainly at the nose of this person.

Figure 6 is another example, Japanese Ex-Prime Minister Hashimoto. In this case, the gallery looks mainly at around the mouth and chin.

4. Caricature Generation Controlled by Gaze Direction

In order to realize the interactive caricaturing system, we introduced the deformation process by using an eye camera which is mounted on the head of the gallery. In this system, the facial parts which a gallery mainly focuses on are selectively deformed. It is expected that this system could directly reflect the visual characteristic information of the gallery onto the caricaturing process. Some works are shown in Fig.7. Here, the values $b_i (i = 1, 2, ..., N)$ are determined based on the analysis of Eq.'s (9)–(11). Since the gallery looked mainly at the nose and cheek of this person, in Fig.7(a), coefficient b_i around the nose and cheeks became large, and for other facial parts it was decided that $b_{i(others)} = 0.4 \sim 0.6$. In the same way in Fig.7(b), since the gallery looked mainly at around the mouth and chin, it was decided that b_i around the mouth and chin became about $1.0 \sim 1.2$ and that $b_{i(others)} = 0.4 \sim 0.6$ for other facial parts by Eq.'s(9)–(11) ($\beta = 0.8, \alpha = 1.0, b_{offset} = 0.4$).

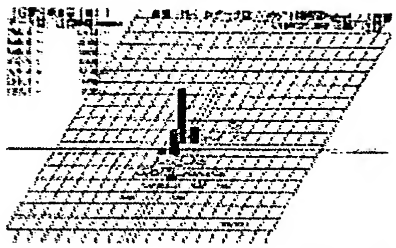
Figure 8 and Fig.9 show other results for comparation. In these figures, the feedback value $K = func_G(P, S)$ from the gallery to the system in Eq.'s(2),(3) is not considered (=0). If there is no feedback information from the gallery, the system (caricaturist) could not decide the proper values b_C, b_G in principle. So that, in these figures, some vague values, 0.00, 0.25, 0.50, 1.00, are given independently to the input face P. Consequently, some unexpected facial parts which are too much deformed or not too much deformed appeared in these caricatures.

5. Conclusions

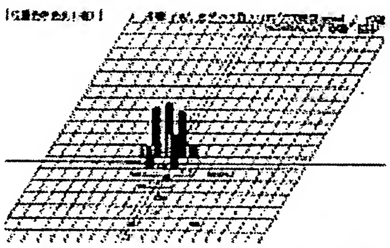
In this paper, (1) a mechanism to extract visual characteristics by using an eye camera, and (2) an interactive caricature generation mechanism controlled by the gaze distribution were proposed. And the effectivity of this method was experimentally demonstrated. It was experimentally known that the proposed method improves the result compared with the conventional method where coefficients $b_i (i = 1, 2, ..., N)$ of all facial parts (points) are constant. This result shows the possibility (a)to extract facial features by using an eye camera, and (b)to represent the gallery's visual characteristics onto facial caricaturing process.

In the experiment, although a static image is used for feature extraction, it is easily noticed and encouraging that it is better for the better facial extraction to utilize the continuous or motion images. To examine how to display these motion images to the gallery and how to analyze the gaze distribution are our future works.

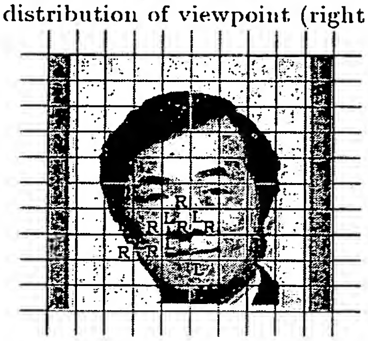
(a) original image



(b) distribution of viewpoint (left eye)

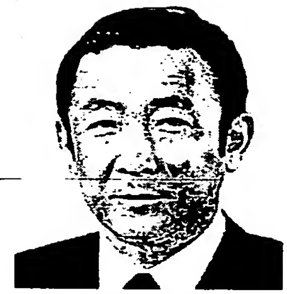


(c) distribution of viewpoint (right eye)

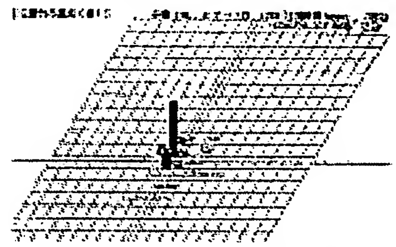


(d) overlaid result

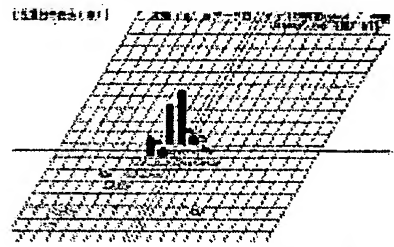
Figure 5. An experimental result of the distribution of viewpoint (for Mr.Naoto Kan).



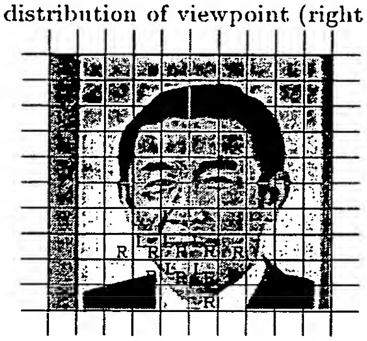
(a) original image



(b) distribution of viewpoint (left eye)

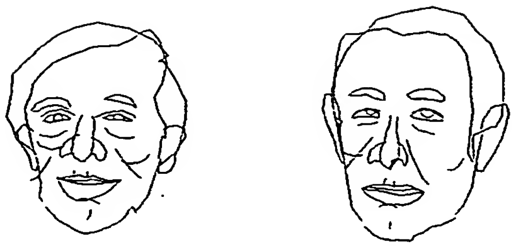


(c) distribution of viewpoint (right eye)



(d) overlaid result

Figure 6. An experimental result of the distribution of viewpoint (for Mr.Ryutaro Hashimoto).



(a) (b)

Figure 7. The caricatures controlled by the interactive system, the degree of interest $b_i(i = 1, 2, ..., N)$ are not constant but selectively distributed, (a) Naoto Kan (left) and (b) Ryutaro Hashimoto (right).

Acknowledgments

This paper was partially supported by Grant-in-Aid for General Scientific Research, IMS HUTOP Research Promotion, and High-Tech.Research Center Promotion.

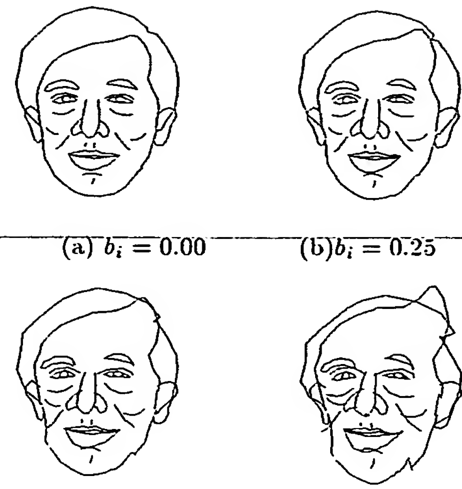
References

[1] For example, "Deformed Caricaturing Gallery," http://www. asahi-net.or.jp/ kq2y-hmgc/index.htm

[2] Brennan,S.E., "Degree of Master of Science in Vision Studies at MIT(Sep.1982).

[3] Dewdney,A.K.: "Computer recreation :computer caricature", *Science*, pp.160-165 (Dec.1986) (in Japanese).

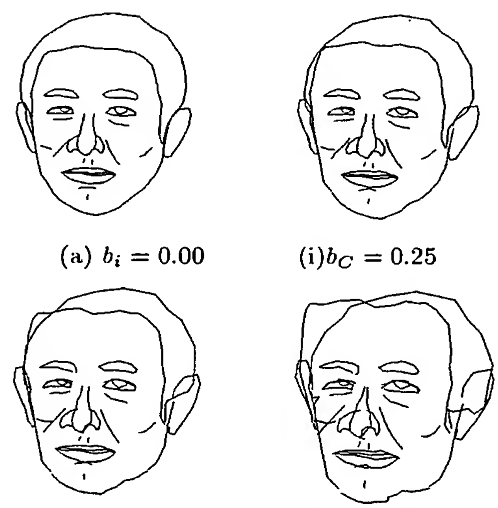
[4] Murakami,K., Takai,M and Koshimizu,H., "On Autonomous Control Method for Facial Caricature Generation Based on a Model of Visual Illusion – Experimental Modeling of Visual Illusion –", *Int'l Conf on Advanced Intelligent Mechatronics(AIM'97)*, CD-ROM Proceedings(June 1997).



(a) $b_i = 0.00$ (b)$b_i = 0.25$

(c) $b_i = 0.50$ (b)$b_i = 1.00$

Figure 8. Examples of caricatures (Naoto Kan), the deformation weight $b_i(i = 1, 2, ..., N)$ is constant.



(a) $b_i = 0.00$ (i)$b_C = 0.25$

(c) $b_i = 0.50$ (i)$b_C = 1.00$

Figure 9. Examples of caricatures (Ryutaro Hashimoto), the deformation weight $b_i(i = 1, 2, ..., N)$ is constant.

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: ______________________

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox